Jet Propulsion Laboratory
California Institute of Technology
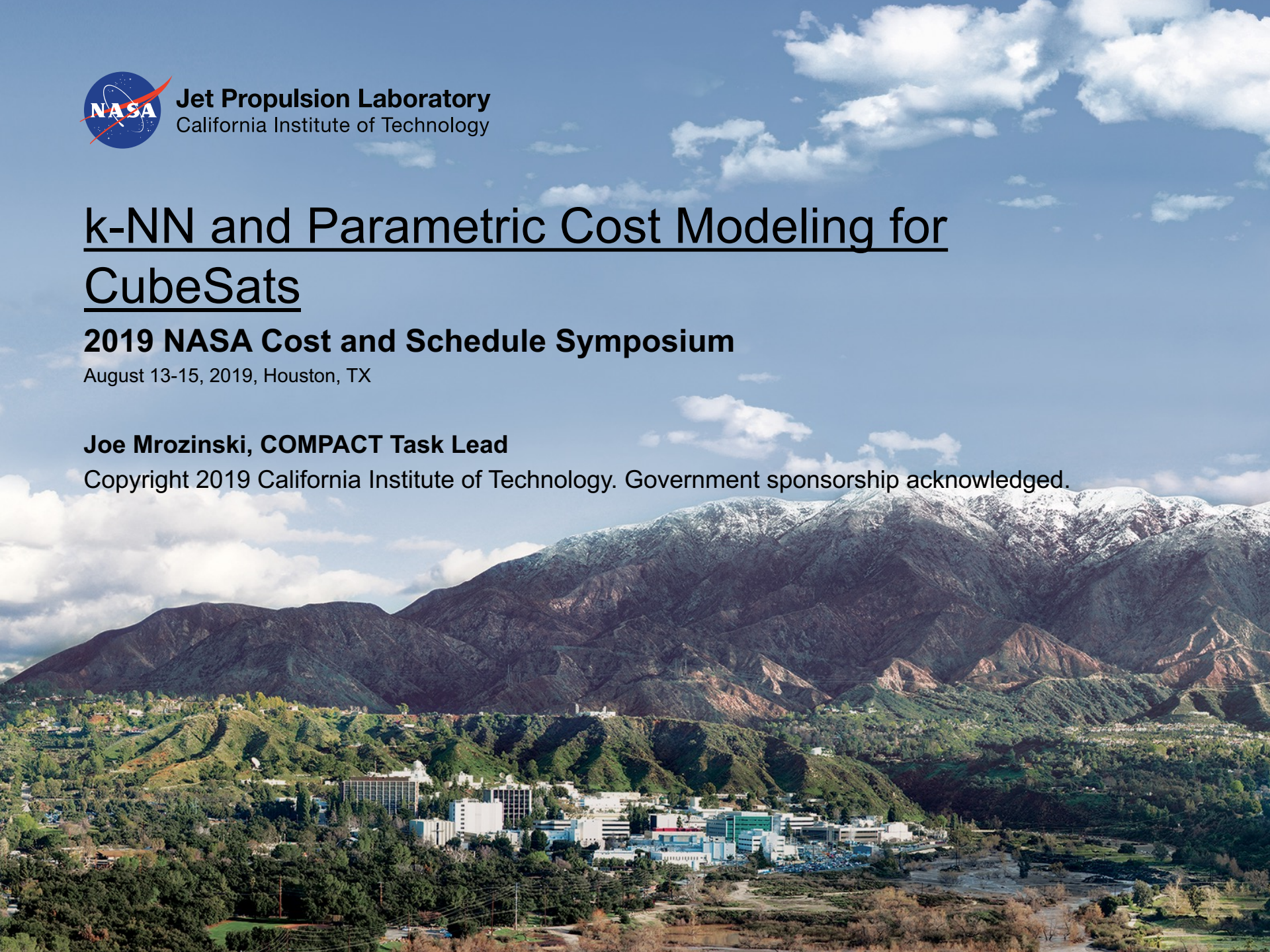
k-NN and Parametric Cost Modeling for CubeSats

2019 NASA Cost and Schedule Symposium

August 13-15, 2019, Houston, TX

Joe Mrozinski, COMPACT Task Lead
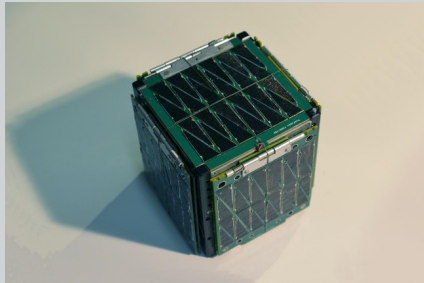
COMPACT Modeling Team

- **Mike DiNicola**
- **Melissa Hooke**
- Elinor Huntington
- Alex Lumnah
- **Joe Mrozinski**
- **Michael Saing**

What is a CubeSat?

CubeSat = Extremely small (i.e. Nanosat scale 1-10kg) spacecraft of standard *dimensions* that hitchhikes to space with a traditional spacecraft.

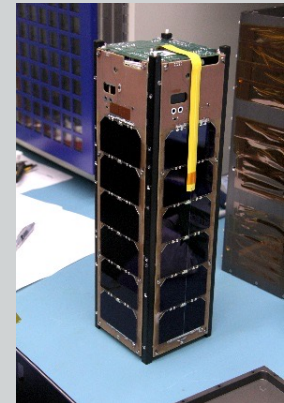- Standard Form Factors: how many "U's" is your Cubesat? 1 to 6 U's:
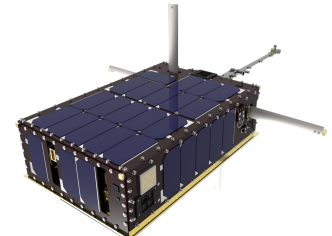
A "1U" Cubesat is roughly 10x10x10 cm
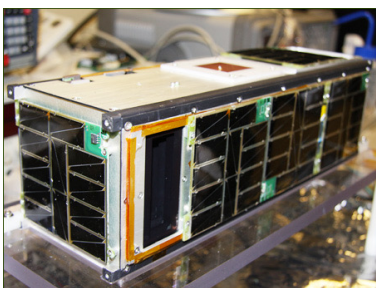

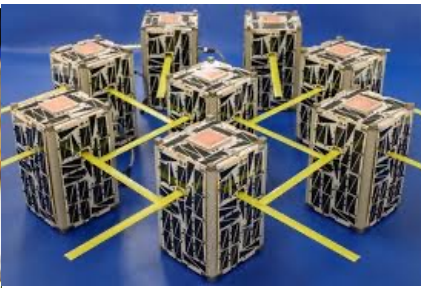A "2U" Cubesat is twice as big

3U:

6U!

Last Year…
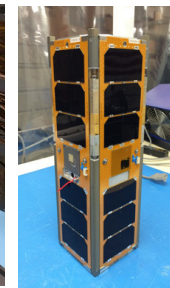
(3 Slide recap)

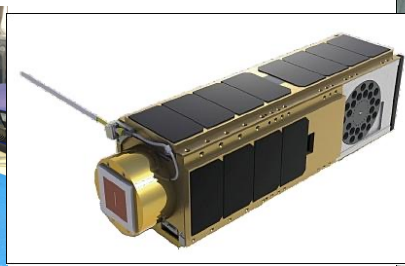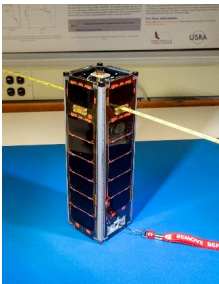# COMPACT has collected and normalized data for
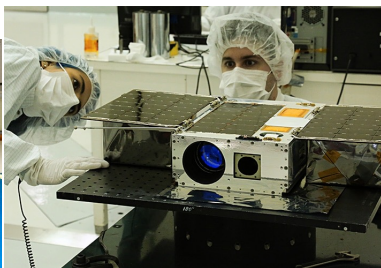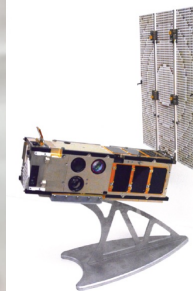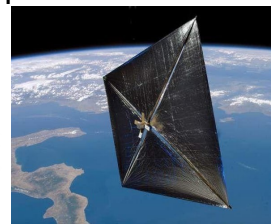
25 CubeSats
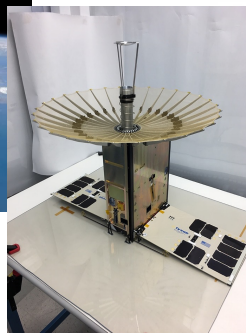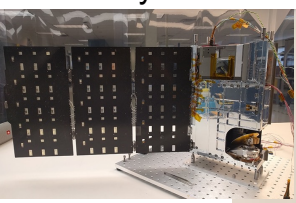
CINEMA

EDSN

GRIFEX

LMRST
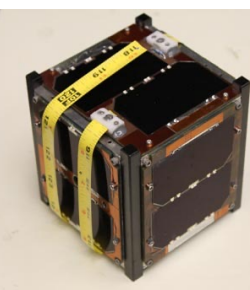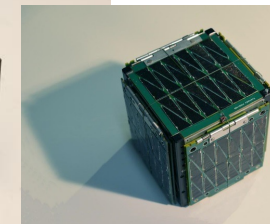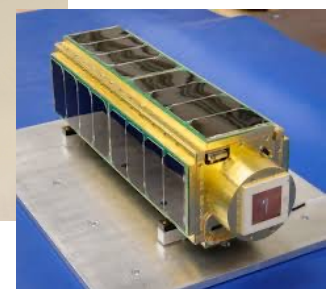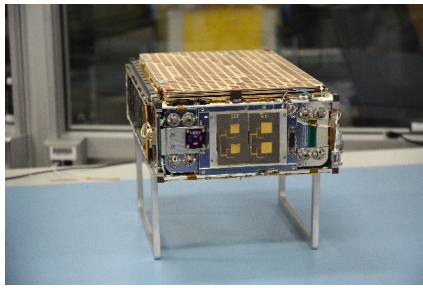
O_OREOS

KickSat

Firefly

ASTERIA

CSUNSAT-1

ISARA

NEA Scout

RainCube

IPEX

TEMPEST-D

M-Cubed 2

SkyCube
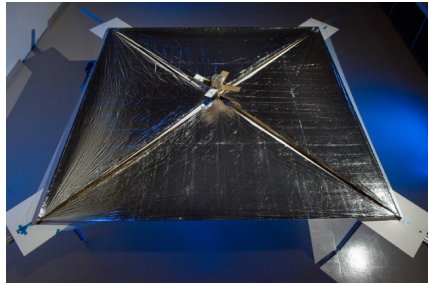
SporeSat-1

PSSC-2

NanoSail-D
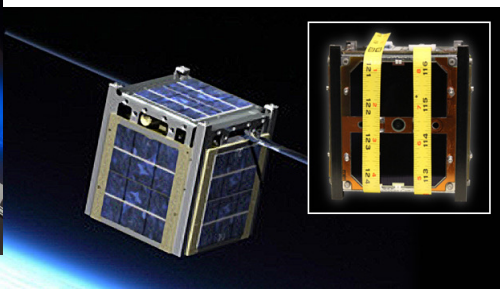
M-Cubed
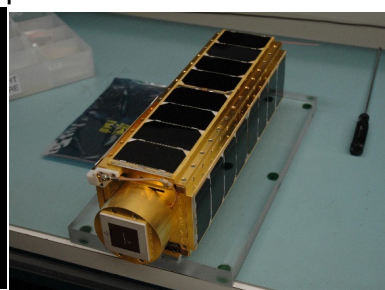
COMPACT
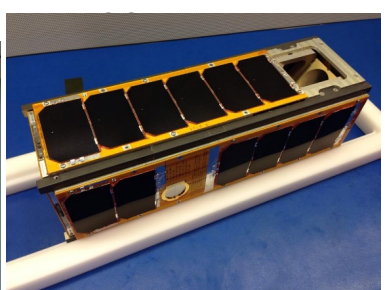
PharmaSat
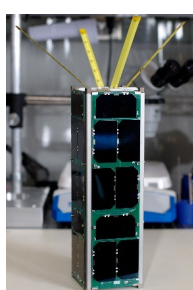
RACE

RAX 1

Key CubeSat Data
i.e. parameters likely to drive cost

CubeSat	Launch Date (Actual or Planned)	Mission Type	Developer Type	# U's	Mass (kg)	Power (W)	Development Schedule (B/C/D)	Design Life (months)
ASTERIA	8/14/2017	Science	JPL	6	11	20	28	3
CINEMA (1)	9/13/2012	Science	University	3	3.15	2.9	44	12
CSUNSat-1	4/18/2017	Educational	University	2	2	4		
DHFR	8/26/2017	Tech Demo	JPL	3	5.03	10		3
EDSN	11/3/2015	Tech Demo	Civil	1.5	2	1	10	24
Firefly (1)	11/20/2013	Science	Civil	3	3.51	3.62	36	3
GRIFEX	1/31/2015	Tech Demo	JPL	3	4			
ISARA	11/10/2017	Tech Demo	JPL	3	5	56	48	5
KickSat (1)	4/18/2014	Tech Demo	University	3	6		4	24
LMRST	10/8/2015	Tech Demo	JPL	3	4.6	8		
MarCO	5/5/2018	Tech Demo	JPL	6	12.7	64	21	6.5
M-Cubed	10/28/2011	Tech Demo	University	1	1	1.2	30	
M-Cubed2	12/5/2013	Tech Demo	University	1	1	1.2		
NanoSail-D (2)	11/20/2010	Tech Demo	Civil	3	4			4
NEA Scout	7/1/2018	Tech Demo	JPL and MSFC	6	12.3	50		
O/OREOS	5/19/2009	Science	Civil	3	5.2		12	18
PharmaSat (1)	5/19/2009	Science	Civil	3	5			
PolySat (CP8) "IPEX"	12/5/2013	Tech Demo	University	1	1	1.5	24	6
PSSC-2	7/10/2011	Tech Demo	Civil	2	3.7	5	6	
RACE	10/28/2014	Tech Demo	JPL	3	5	1.5		
RainCube	5/20/2018	Tech Demo	JPL	6	12	35	17	2
RAX 1 (USA 218)	11/20/2010	Science	University	3	3	8		12
SkyCube	1/9/2014	Educational	Commercial	1	1.3	4	24	3
SporeSat-1	4/18/2014	Tech Demo	Civil	3	5.2		36	2
Tempest-D	2/1/2018	Tech Demo	JPL	6	14	21	21	3

Present Day…

Cost Estimating Approaches

CubeSat Cost Estimating Approaches

Using the data collected in the previous effort, we examined 2 cost estimation approaches:

1. K-Nearest Neighbors (k-NN)
2. Parametric Cost Modeling

CubeSat Cost Estimating Approaches

Using the data collected in the previous effort, we examined 2 cost estimation approaches:

1. K-Nearest Neighbors (k-NN) **– Completed**
2. Parametric Cost Modeling **– Sneak Peak**

K-Nearest Neighbors

- **Created a K-Nearest Neighbors Analogy-drive cost model for CubeSats utilizing the framework developed by the ASCoT Team.**
- **This model is accessed via the Web and will soon be available via ONCE.**

K-Nearest Neighbors

KNN is a simple form of analogy cost estimation. Here's how it works:

User Inputs	Find K "closest" observations in historical data	Calculate weighted average of those K costs	**Algorithm Outputs**
• Mass (kg) • # of U's • # of Spacecraft • Developer type (university, commercial, civil, JPL)			• Names of K most similar CubeSats • Cost estimate for set of inputs
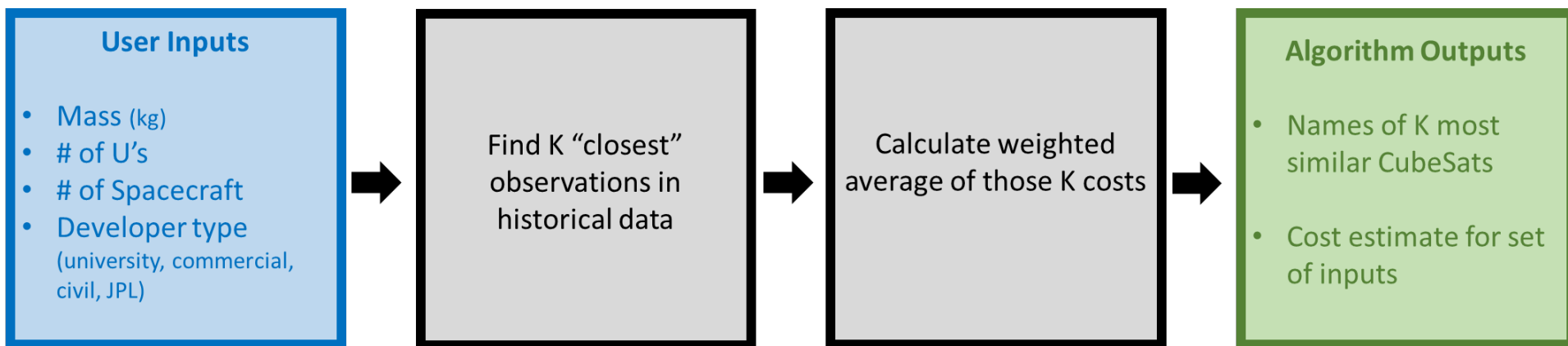
"Closest" here is determined by Euclidean distance between points. Now, the only thing left to do is to choose the number of neighbors, K.

K-Nearest Neighbors
K is chosen by leave-one-out cross validation

1. *Remove one observation, the "test mission", from the data*
2. *Build KNN model on the remaining observations or "training missions"*
3. *Predict cost for the test mission using weighted average of k neighbors*
4. *Compute MRE statistic for the test mission*
5. *Repeat Steps 2-5 for all n missions*
6. *Compute MRE summary statistics given n errors from Steps 6*
7. *Repeat 1-7 for each new k and choose the k that minimizes the most parameters in Step 7*

K	25th	50th	75th	Mean
1	0.29	0.63	1.05	1.87
2	0.16	0.45	0.89	1.66
3	0.22	0.42	0.89	1.75
4	0.27	0.60	0.96	1.85
5	0.21	0.57	0.96	1.96

MRE Summary for Different K Values

K-Nearest Neighbors

The MRE for each mission is calculated based on inputting each mission and outputting a KNN prediction after removing that CubeSat from the training data. The results are compared to MREs from using the mean and the median of the training data which is considered to be a naïve estimate good for model assessment.

KNN MRE Performance vs. Naive Estimate
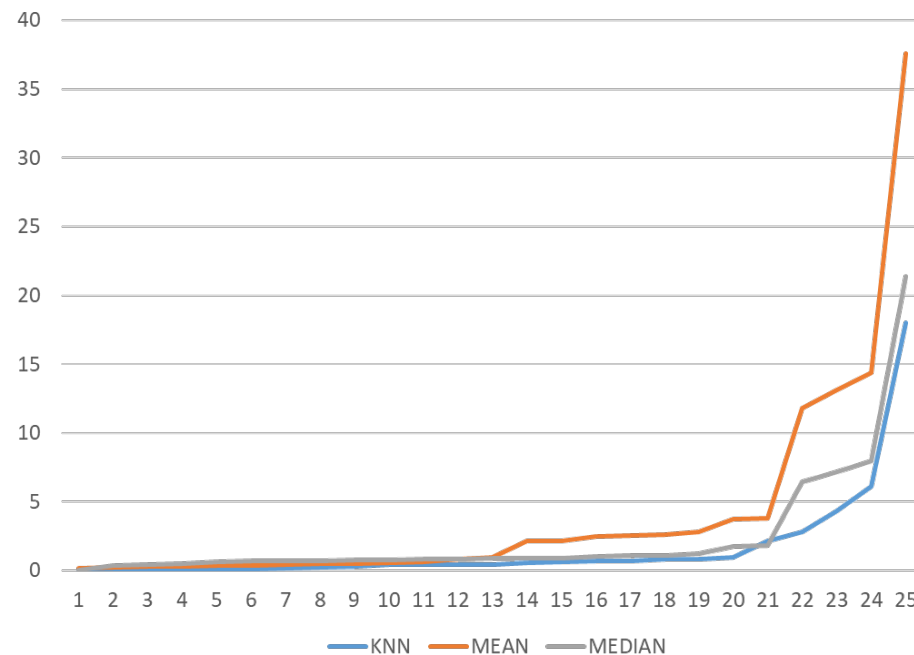
KNN MRE Performance vs. Naive Median Estimate

...eb Tool

Jet Propulsion Laboratory
California Institute of Technology

k-NN and Parametric Cost Modeling for CubeSats

2019 NASA Cost and Schedule Symposium
August 13-15, 2019, Houston, TX

Joe Mrozinski, COMPACT Task Lead
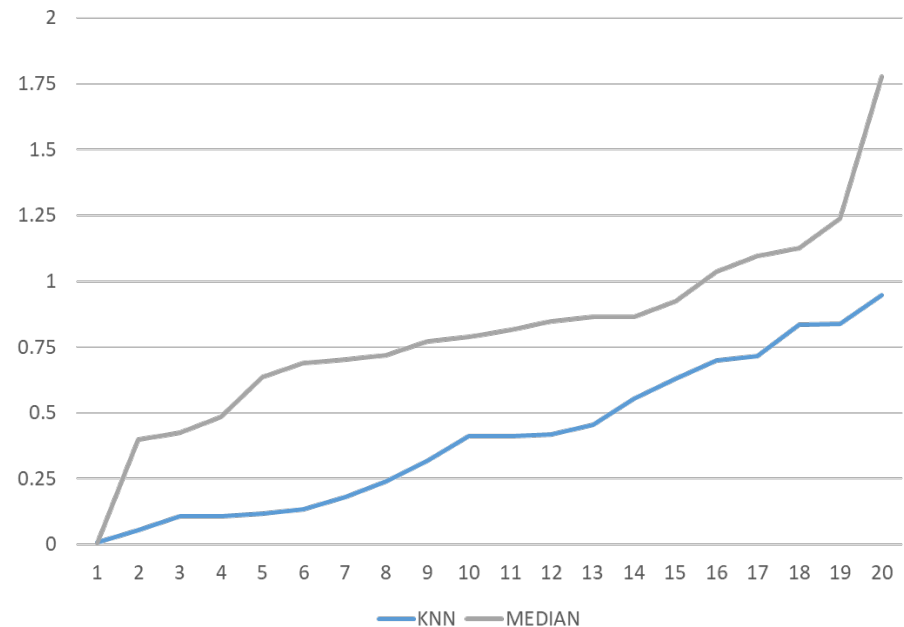Copyright 2019 California Institute of Technology. Government sponsorship acknowledged.

...a.gov/knn/compact/

...ompact Knn Cost Estimator

...he K-Nearest Neighbor Regression Algorithm is a simple non-parametric method used to ...stimate the total cost to develop a CubeSat mission based on previous missions. Using a ...andful of inputs, the model assigns a distance metric that ranks each mission in order of ...milarity to the estimate mission.

Create New Estimate

Estimate Name	JPL / NASA Developed
New CubeSat	Yes

Mass	U's
1.00	1

Number of Spacecraft
1

[Import Inputs (CSV File)] [Create Estimate]
[Export Inputs (CSV File)]
[Choose File] No file chosen

Current Estimate

Estimated Cost: 1333 k$

KNN Results Summary

Neighbors	Cost (k$)	Distance
PSSC-2	1200.0	0.29
Firefly (1)	2578.0	0.44
NanoSail-D (2)	250.0	0.46

Estimate

The cost information contained in this document is of a budgetary and planning nature and is intended for informational purposes only. It does not constitute a commitment on the part of JPL and/or Caltech.

K-Nearest Neighbors Web Tool

Create New Estimate

Estimate Name
New CubeSat

JPL / NASA Developed
Yes ▾

Mass
1.00

U's
1

Number of Spacecraft
1

[Import Inputs (CSV File)] [Create Estimate]

[Export Inputs (CSV File)]

[Choose File] No file chosen

Current Estimate

Estimated Cost: 1333 k$

KNN Results Summary

| Neighbors	Cost (k$)	Distance
PSSC-2	1200.0	0.29
Firefly (1)	2578.0	0.44
NanoSail-D (2)	250.0	0.46

K-Nearest Neighbors Web Tool

Results: Estimated Cost Compared to Nearest Neighbors and all other Missions

This chart displays the variation in CubeSat Cost for all missions. The X indicates the effort of the estimate mission.

Parameter Distance to Nearest Neighbors

Cost
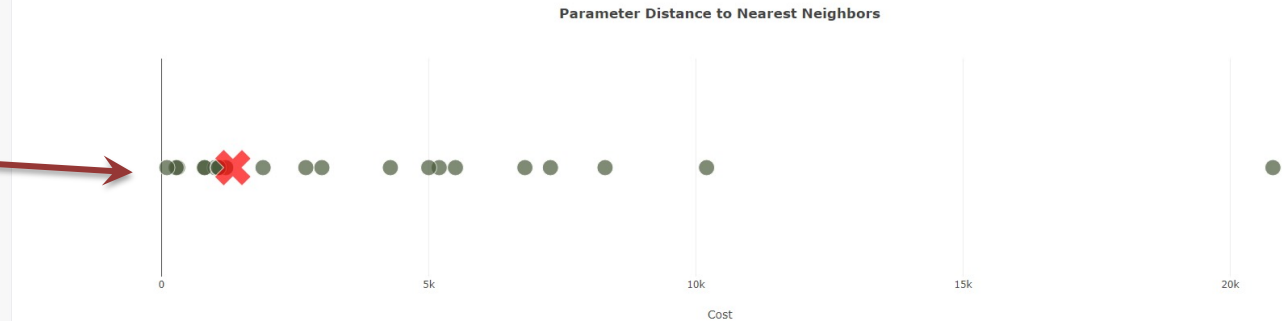
This chart displays the Euclidean distance from the estimate mission for all missions in the data set.

Parameter Distance to Nearest Neighbors

Results by Euclidian Distance

Euclidean Distance Across All Parameters
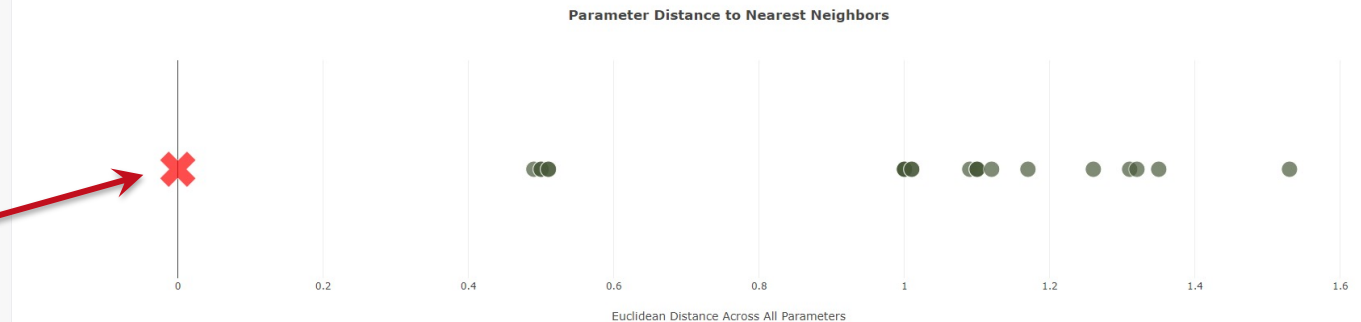
Parametric Models – Sneak Peak

Parametric Models – Sneak Peak

- Apply stepwise and best-subsets regression methods to identify potential CubeSat parametric cost models.

- Utilize ANOVA, standard significance tests and R^2 to identify potential cost drivers and compare/select best models.
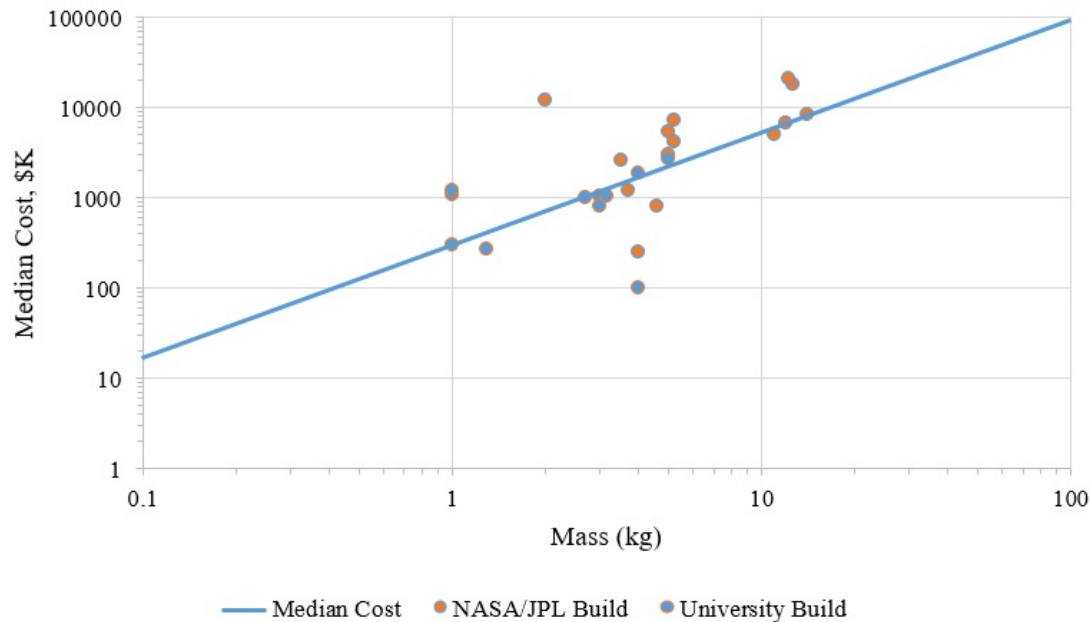
Draft Preliminary Beta Candidate Model #1

AKA "Not ready for use in Proposal Development/Evaluation"

Draft Preliminary Beta Candidate Model #1

AKA "Not ready for use in Proposal Development/Evaluation"
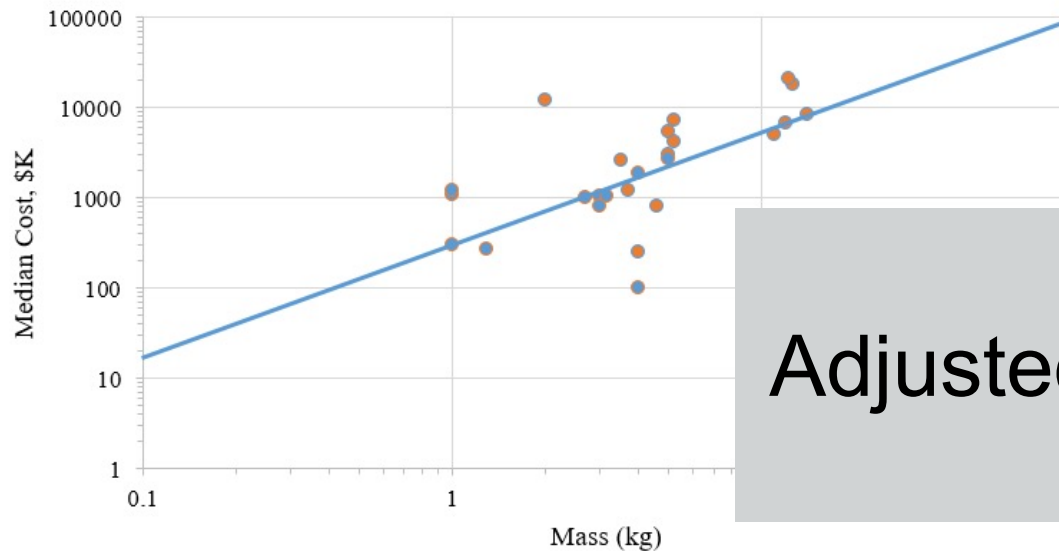

Developer Type & Costs Vs Mass

$$Cost = e^{6.0759} e^{1.0987 \times \ln(Mass)}$$

$$= 435 \times Mass^{1.0987}$$

Draft Preliminary Beta Candidate Model #1

AKA "Not ready for use in Proposal Development/Evaluation"

Developer Type & Costs Vs Mass

Adjusted R^2: 36%
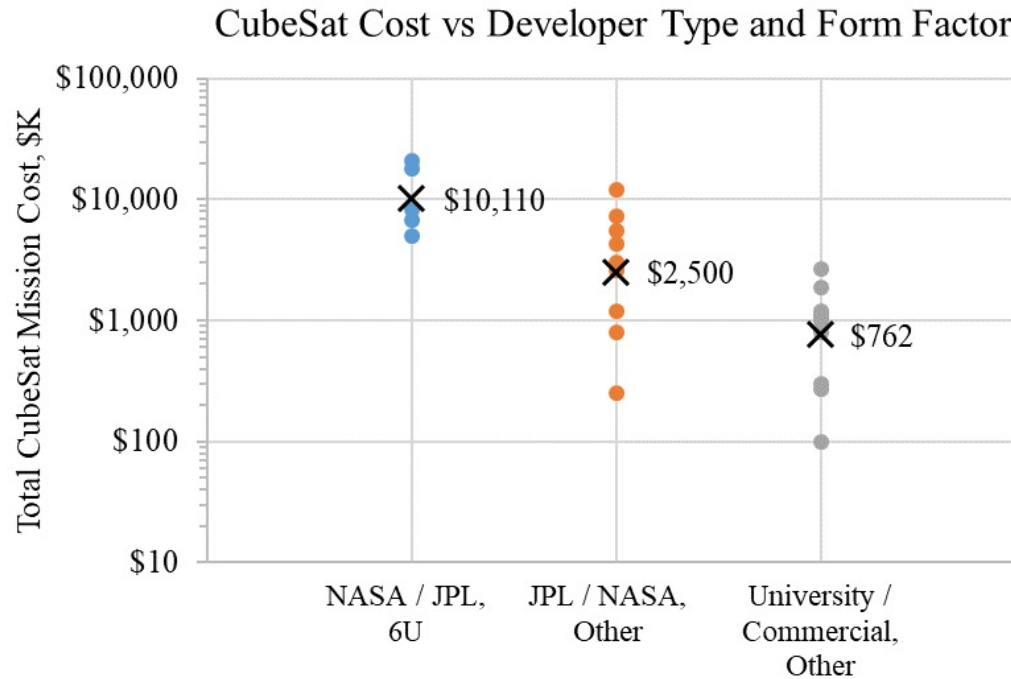
$$\text{Cost} = e^{6.0759} e^{1.0987 \times \ln(\text{Mass})}$$

$$= 435 \times \text{Mass}^{1.0987}$$

Draft Preliminary Beta Candidate Model #2

AKA "Not ready for use in Proposal Development/Evaluation"

Draft Preliminary Beta Candidate Model #2

AKA "Not ready for use in Proposal Development/Evaluation"

CubeSat Cost vs Developer Type and Form Factor

$$
\text{Cost} = \begin{cases} 10,110 = e^{7.8241}e^{1.3972} & \text{if 6U \& DevGrp is NASA/JPL} \\ 2,500 = e^{7.8241} & \text{if <6U \& DevGrp is NASA/JPL} \\ 762 = e^{7.8241}e^{-1.1882} & \text{if <6U \& DevGrp is Univ/Comm*} \end{cases}
$$

COMPACT 24 **jpl.nasa.gov**

*No data collected yet for <6U CubeSats developed by Univ/Comm

Draft Preliminary Beta Candidate Model #2

AKA "Not ready for use in Proposal Development/Evaluation"

CubeSat Cost vs Developer Type and Form Factor

Adjusted R^2: 47%
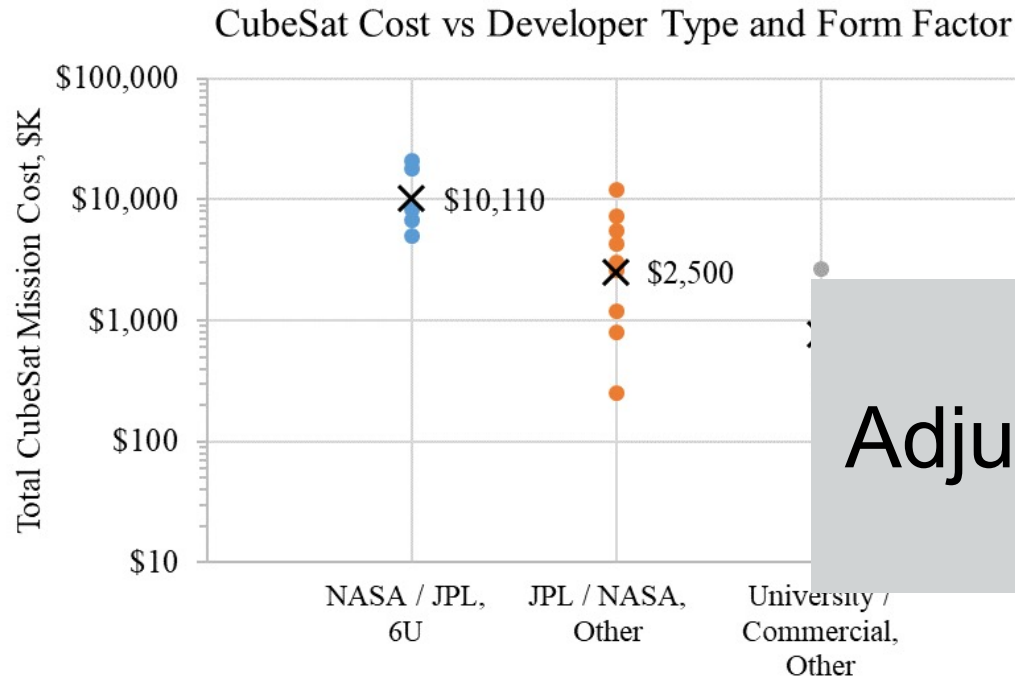
$$\text{Cost} = \begin{cases} 10,110 = e^{7.8241}e^{1.3972} & \text{if 6U \& DevGrp is NASA/JPL} \\ 2,500 = e^{7.8241} & \text{if <6U \& DevGrp is NASA/JPL} \\ 762 = e^{7.8241}e^{-1.1882} & \text{if <6U \& DevGrp is Univ/Comm*} \end{cases}$$

*No data collected yet for <6U CubeSats developed by Univ/Comm

Concluding Thoughts

- The COMPACT Team has:
  - Delivered a normalized database of flown CubeSats missions.
  - Utilized that database to create a Analogy-Based Cost Estimation Tool using k-NN
  - Created Candidate Parametric Models which show promise, but required further data collection and refinement before delivering CERs for proposal/evaluation use.

Acknowledgements and Thanks

- NASA OCFO for Sponsoring this task

- Aerospace Corporation for their aid in our initial data collection/normalization phase

- All of the CubeSat project personnel who provided data.

Questions?

jpl.nasa.gov